OIPE IAP35
NOV 14 2008
PATENT & TRADEMARK OFFICE

IF AF

Docket No. 10020708-1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of:

Applicants	:	Zohar Yakhini et al.
Application No.	:	10/825,893
Filed	:	April 16, 2008
For	:	METHOD AND SYSTEM FOR NORMALIZATION OF MICROARRAY DATA

Examiner	:	NEGIN, Russell Scott
Art Unit	:	1631
Docket No.	:	10020708-1
Date	:	November 11, 2008

APPEAL BRIEF

Mail Stop: Appeal Briefs – Patents
Commissioner of Patents and Trademarks
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This appeal is from the decision of the Examiner, in an Office Action mailed April 30, 2008, finally rejecting claims 1-12 and 14-28.

REAL PARTY IN INTEREST

Agilent Technologies is the Assignee of the present patent application. Agilent Technologies, Inc., is a Delaware corporation with headquarters in Santa Clara, California.

RELATED APPEALS AND INTERFERENCES

Applicant's representative has not identified, and does not know of, any other appeals of interferences which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

STATUS OF CLAIMS

Claims 1-12 and 14-28 are pending in the application. Claims 1-12 and 14-28 were finally rejected in the Office Action dated April 30, 2008. Claims 13 and 29 have been canceled. Applicants' appeal the final rejection of claims -12 and 14-28 which are copied in the attached CLAIMS APPENDIX.

STATUS OF AMENDMENTS

No Amendment After Final is enclosed with this brief. The last Amendment was filed October 2, 2007.

SUMMARY OF CLAIMED SUBJECT MATTER

Independent Claim 1

Claim 1 is directed to a method (page 23, line 10 to page 18, line 18) for selecting a set of normalizing data points (1452, 1454) from *n* data sets (1405-1408), where *n* is at least 3, containing data points having values (1412, 1414, 1416, 1418) and identities (1410`), the method comprising: (1) receiving *n* data sets (page 39, lines 9-19); (2) considering the data points to be distributed in an *n*-dimensional data-point space (Figure 10A; Figure 11A; Figure 12); (3) determining one or more order-preserving sequences (1440-1450; page 17, lines 11-16) of data points within the *n*-dimensional data-point space; (4) selecting, as normalizing data points, data points from the one or more order-preserving sequences (page 27, lines 9-20); and (5) storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets (page 36, line 19 to page 37, line 23).

Dependent Claims 2-12

Claim 2 is directed to the method of claim 1 wherein the one or more order-preserving sequences of data points (page 17, lines 11-16) is a single, longest order-preserving sequence of data points . Claim 3 is directed to the method of claim 1 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights (page 17, lines 16-22). Claim 4 is directed to the method of claim 1

wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence (page 44, lines 13-16). Claim 5 is directed to the method of claim 1 wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points of lengths within a threshold value of the length of an order-preserving sequence of data points of maximum length (page 26, line 27 to page 27, line 26). Claim 6 is directed to the method of claim 1 wherein the data points within n data sets are associated with weights and wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points with sums of weights within a threshold value of the sum of weights of an order-preserving sequence of data points with a greatest sum of weights (page 26, line 27 to page 27, line 26). Claim 7 is directed to the method of claim 1 wherein considering the data points to be distributed in an n-dimensional data-point space further includes, for each data point, considering the data point to have a value in each of n-dimensions, the value of a data-point in an ith dimension equal to the value of the data point in an ith data set, where $1 \leq i \leq n$ ((1412, 1414, 1416, 1418). Claim 8 is directed to the method of claim 1 wherein determining an order-preserving sequence of data points within the n-dimensional data-point space further includes (page 23, line 10 to page 18, line 18): (1) for each currently considered dimension, ordering the data points with respect to the currently considered dimension; traversing the ordered data points in a first direction, determining a metric corresponding to a maximum subsequence for each data point in the first direction; and traversing the ordered data points in a second direction, determining a metric corresponding to a maximum subsequence for each data point in the second direction; (2) summing the determined metrics for each data point in each dimension to produce a metric sum for each data point; and (3) selecting as belonging to the maximum order-preserving sequence of data points those data points having a greatest metric sum. Claim 9 is directed to the method of claim 8 wherein selecting, as normalizing data points, data points from the order-preserving sequence further includes selecting data points with a metric sum greater than a threshold value (page 26, line 27 to page 27, line 26). Claim 10 is directed to the method of claim 8 wherein selecting, as normalizing data points, data points from the one or more order-preserving sequences further includes selecting data points of a single order-preserving sequence (page 27, lines 9-20). Claim 11 is directed to the method of claim 8 wherein selecting, as normalizing data points, data points from the one or more order-

preserving sequences further includes selecting data points that most evenly partition the data points into subsets of data points (page 27, lines 9-20). Claim 12 is directed to computer instructions (page 33, line 20 to page 44, line 3) stored in a computer readable medium that implement the method of claim 1.

Independent Claim 14

Claim 14 is directed to a system for selecting (page 23, line 10 to page 18, line 18) a set of normalizing data points (1452, 1454) from *n* data sets (1405-1408), where *n* is at least 3, containing data points having values (1412, 1414, 1416, 1418) and identities (1410), the system comprising: (1) a processor; (2) a memory; and (3) computer instructions (page 33, line 20 to page 44, line 3) that select the set of normalizing data points from *n* data sets by receiving *n* data sets (page 39, lines 9-19), considering the data points to be distributed in an *n*-dimensional data-point space (Figure 10A; Figure 11A; Figure 12), determining one or more order-preserving sequences (1440-1450; page 17, lines 11-16) of data points within the *n*-dimensional data-point space, and selecting, as normalizing data points, data points from the one or more order-preserving sequences; and storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets (page 27, lines 9-20).

Dependent Claims 15-19

Claim 15 is directed to the system of claim 14 wherein the one or more order-preserving sequences of data points (page 17, lines 11-16) is a single, longest order-preserving sequence of data points. Claim 16 is directed to the system of claim 14 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights (page 17, lines 16-22). Claim 17 is directed to the system of claim 14 wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence (page 44, lines 13-16). Claim 18 is directed to the system of claim 14 wherein the one or more order-preserving sequences of data points are order-preserving sequence of data points within a threshold value of an order-preserving sequences of data points of maximum length (page 26, line 27 to page 27, line 26). Claim 19 is directed to the

system of claim 14 wherein the one or more order-preserving sequences of data points are order-preserving sequence of data points within a threshold value of an order-preserving sequences of data points with a greatest sum of weights (page 26, line 27 to page 27, line 26).

Independent Claim 20

Claim 20 is directed to the method for selecting a set of normalizing data points (1452, 1454) from *n* data sets (1405-1408), where *n* is at least 4 and even, containing data points having values and identities, the method comprising: (1) receiving *n* data sets (page 39, lines 9-19); (2) considering the data points to be distributed in $\frac{n}{2}$ 2-dimensional data-point spaces (Figure 10A; Figure 11A; Figure 12); (3) determining one or more order-preserving sequences of data points (1440-1450; page 17, lines 11-16) for each of the $\frac{n}{2}$ 2-dimensional data-point spaces, (4) selecting, as normalizing data points, data points from the order-preserving sequences (page 27, lines 9-20); and (5) storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets (page 36, line 19 to page 37, line 23).

Dependent Claims 21-28

Claim 21 is directed to the method of claim 20 wherein the one or more order-preserving sequences of data points (page 17, lines 11-16) is a single, longest order-preserving sequence of data points. Claim 22 is directed to the method of claim 20 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights (page 17, lines 16-22). Claim 23 is directed to the method of claim 20 wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence (page 44, lines 13-16). Claim 24 is directed to the method of claim 20 wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points within a threshold value of an order-preserving sequence of data points of maximum length (page 26, line 27 to page 27, line 26). Claim 25 is directed to the method of claim 20 wherein the data points within *n* data sets are associated with

weights and wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points with sums of weights within a threshold value of the sum of weights of an order-preserving sequence of data points with a greatest sum of weights (page 26, line 27 to page 27, line 26). Claim 26 is directed to the method of claim 20 wherein determining an order-preserving sequence of data points within a 2-dimensional data-point space further includes (page 23, line 10 to page 18, line 18): (1) for each currently considered dimension, ordering the data points with respect to the currently considered dimension; traversing the ordered data points in a first direction, determining a metric corresponding to a maximum subsequence for each data point in the first direction; and traversing the ordered data points in a second direction, determining a metric corresponding to a maximum subsequence for each data point in the second direction; (2) summing the determined metrics for each data point in each dimension to produce a metric sum for each data point; and (3) selecting as belonging to the maximum order-preserving sequence of data points those data points having a greatest metric sum. Claim 27 is directed to the method of claim 20 wherein selecting, as normalizing data points, data points from the one or more order-preserving sequences further includes selecting data points which occur in the one or order-preserving sequences computed for greater than a threshold fraction of the $\frac{n}{2}$ 2-dimensional data-point spaces (page 26, line 27 to page 27, line 26). Claim 28 is directed to computer instructions (page 33, line 20 to page 44, line 3) stored in a computer readable medium that implement the method of claim 20.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

1. The rejection of claims 1-12 and 14-28 under 35 U.S.C. §101.

2. The rejection of claims 1, 4, 7, 12, 14, 17, 20, 23, and 28 under 35 U.S.C. §103(a) as being unpatentable over Larson et al., Calculus with Analytic Geometry, 1990, D.C. Heath and Company; Lexington, Massachusetts; Section 14.1, pages 27-33, 785-795, and page 840.

ARGUMENTS

Claims 1-12 and 14-28 are pending in the current application. In an office action dated April 30, 2008 ("Office Action"), the Examiner finally rejected claims 1-12 and 14-28 under 35 U.S.C. §101 as being directed to non-statutory subject matter and rejected claims 1, 4, 7, 12, 14, 17, 20, 23, and 28 under 35 U.S.C. §103(a) as being unpatentable over Larson et al., Calculus with Analytic Geometry, 1990, D.C. Heath and Company; Lexington, Massachusetts; Section 14.1, pages 27-33, 785-795, and page 840 ("Larson"). Appellants' representative respectfully traverses these rejections.

ISSUE 1

1. The rejection of claims 1-12 and 14-28 under 35 U.S.C. §101.

In rejecting claims 1-12 and 14-28 under 35 U.S.C. §101, the Examiner cites 35 U.S.C. §101 and a portion of the "Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility" ("Guidelines") It is Appellants' representative's understanding that the Guidelines have been superseded by new material introduced in the M.P.E.P. Indeed, the portion of the Guidelines quoted by the Examiner appears in M.P.E.P. §2106(IV)(C)(1)(2) Practical Application That Produces A Useful, Concrete, And Tangible Result.

After quoting 35 U.S.C. §101 and a portion of the Guidelines, now found in the above-referenced section of the M.P.E.P., the Examiner states:

> The instant claims are drawn to computational means for selecting a set of normalizing data points. However, as claimed, the method does not produce a tangible result. For example, the method as claimed may take place entirely within the confines of a computer or a human mind without any communication to the outside world and without using or making available for use, the results of the computation. In this instance, it is possible that the memory is only accessible by other computer memories, in which case the resultant data would not be accessible to a user. Thus, the instant methods of the claims do not produce any tangible result.
>
> Likewise, while claims 14-19 recite a SYSTEM for selecting a set of normalizing data points, and the processor and memory are tangible, the computer instructions execute a method that does not necessarily produce a tangible result (i.e. storing data of the selected normalizing points in a computer memory may be only accessible by other computer memories and not necessarily a user.).
>
> Additionally claims 12 and 28 are drawn to computer instructions stored on a computer readable medium used to implement the methods described above. Since computer readable media are not defined in the specification, they are interpreted to encompass carrier waves, which are per se, not statutory.

The Examiner appears to conclude that, should it be possible that memories may be only accessible by other computer memories, the resultant data would not accessible to a user, and therefore the claims do not produce a tangible result. Appellants' representative cannot see any rationale or logic behind this conclusion. Nothing in the quoted portion of the Guidelines discusses any requirement that a tangible result be accessible to a user. Furthermore, the statement that "it is possible that the memory is only accessible by other computer memories" makes little sense with respect to computer architecture and hardware. Memories do not access other memories. Memories are accessed by other computer components under the control of logic circuits or routines running on processors. Memories store data, written to memories and read from memories by access operations carried out by routines executing on a processor. It is therefore not possible, according to the clear language of claim 1, "that the memory is only accessible by other computer memories." The statement makes even less sense when considering that the final element of claim 1 cites "storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets." Those familiar with computer science and computer hardware well understand that memories do not normalize data sets or execute any other type of method step. Memories store data, written to memories and read from memories by access operations carried out by routines executing on a processor. It is therefore not possible, according to the clear language of claim 1, "that the memory is only accessible by other computer memories," since the stored normalizing data points are subsequently used for normalization by executing routines and programs. By the Examiner's reasoning, operating systems and much of the internal circuitry and mechanical components of a computer system would be non-statutory subject matter for patenting, since many of these components do not return results to users, but instead essentially affect state transitions within computers in a deterministic fashion, allowing other components and processes to execute. However, it is clear that, in fact, operating systems and much of the internal circuitry and mechanical components of a computer system are indeed quite patentable.

The Examiner does not cite or refer to any statute, rule, or decision for the proposition that a result not provided directly to a user of a system is necessarily intangible. Furthermore, at the beginning of the quoted portion of the Guidelines, it is clear that a basic requirement is that a claimed method or component bring about a useful application. No one familiar with scientific observation and data processing disputes that the selection of

normalizing data points for subsequent normalization of data sets is not a useful application of the claimed method. In fact, a great deal of effort has been undertaken by Appellants to describe the data normalization problem in the current application, beginning on line 26 of page 5 and extending to line 20 of page 9. In the paragraph that begins on line 21 of page 9 of the current application, the need for normalizing data-point determination, to which method and system embodiments of the present invention are directed, is clearly stated. No one familiar with life sciences, scientific data processing, and basic mathematics would doubt the utility of the currently claimed method and system.

The Examiner states that since computer-readable media are not defined in the specification, they are interpreted to encompass carrier waves, which are *per se* not statutory. In essence, the Examiner is arbitrarily defining the term "computer-readable media." However, Appellants' representative does not believe that arbitrary definitions can be used to interpret claim phrases. Furthermore, in the paragraph beginning on line 15 of page 16 of the current application, Appellants discuss handling of data scanned from a microarray. Several times in this paragraph, Appellants refer to "results which are stored in a computer-readable medium, transferred to an intercommunicating entity via electronic signals, printed in a human-readable format, or otherwise made available fur further use." Clearly, Appellants do not intend the phrase "computer-readable media" to encompass transfer via electronic signals, or "carrier waves." The Examiner's arbitrary definition is inconsistent with the use of the term "computer-readable media" in the current application. Furthermore, those with even cursory familiarity with computer science, computer hardware, and electronics well understand that, for data to be encoded in carrier waves, the data must first be encoded in a physical medium, such as an electronic memory. Transmitters, which transmit carrier waves and data encoded with various types of modulation techniques, retrieve the data from physical devices prior to encoding the data in radio-frequency or optical signals. The Examiner's attempt to force the phrase "computer-readable medium" to mean "carrier waves" is therefore inconsistent with the current application, inconsistent with well-understood principles of computing and electronics, and inconsistent with basic principles of claim interpretation.

In Appellants' representative's respectfully offered opinion, the Examiner has failed to provide even rudimentary support in case law, rule, or statute for the Examiner's conclusions that the current claims are non-statutory for failing to produce a tangible result. The quoted passage from the Guidelines appears to state that physical transformation is not a

requirement for subject matter patentability; but that if a claim does not entail the transformation of an article, then the Examiner is to review a claim to determine if the claim provides a practical application that produces a useful, tangible, and concrete result. Those familiar with computer science and electronics well understand that storing normalization points in a computer memory most certainly entails the transformation of an article, namely the electronic state of the computer memory. Furthermore, much of the discussion in the Guidelines appears to be concerned with deciding whether algorithms or disembodied software are patentable. In the case of dependent claims 1, 14, and 20, there is, in Appellants' representative's respectfully offered opinion, absolutely no question that the claimed subject matter does not constitute merely an algorithm or disembodied software. These are methods and systems that produce practical results which are electronically stored and subsequently accessed for carrying out data normalization. Disembodied software and abstract algorithms do not store data in memories, do not select normalizing data points, do not determine order-preserving sequences, and do not carry out any of the other currently claimed steps.

It is a bedrock principle of administrative-law adjudication that decisions need to be based on rules, statutes, and previous legal decisions rather than arbitrary and capricious decisions. In Appellants' representative's respectfully offered opinion, the U.S.P.T.O. is currently rejecting claims under 35 U.S.C. §101 in a very arbitrary manner. In the current case, rejections may be based on conclusions that do not in any way follow from published decisions, rules, or statutes. Certainly the current decision does not follow from the text of 35 U.S.C. §101 or the quoted portion of the Guidelines. In Appellants' representative's respectfully offered opinion, an examiner is obligated to provide a rational basis for rejections, and not merely conclusory statements without support in rule, statute, or legal decisions. Otherwise, it is simply impossible for Appellants to have any hope of drafting claims in a rational matter to obtain predictable results.

ISSUE 2

2. <u>The rejection of claims 1, 4, 7, 12, 14, 17, 20, 23, and 28 under 35 U.S.C. §103(a) as being unpatentable over Larson et al., Calculus with Analytic Geometry, 1990, D.C. Heath and Company; Lexington, Massachusetts; Section 14.1, pages 27-33, 785-795, and page 840</u>.

In the current application, beginning on line 26 of page 5, the fact that multiple

data sets may be obtained from microarray experiments is discussed. However, as discussed in the paragraph beginning on line 28 of page 6, individual data sets within a set of multiple data sets need to be normalized with one another in order for scientifically meaningful conclusions to be drawn from multiple data sets.

Data normalization is discussed in the current application beginning on line 8 of page 7. In the first paragraph of the discussion of data normalization, an example is provided with reference to Figure 8. As discussed in the paragraph beginning on line 3 of page 8 of the current application, two different data sets are shown in Figure 8 to be generally related by a constant intensity-value shift, or translation in the vertical direction. As discussed in the paragraph beginning on line 14 of page 8, and as illustrated in Figure 9, data sets can be normalized to a common intensity scale, so that those features of the microarray that produce signals that differ from one another, apart from the above-mentioned translational shift, can be recognized. Data-set normalization is an extremely well-known problem in all branches of observational sciences. Beginning on line 28 of page 8 of the current application, various ways for carrying out normalization are discussed. However, these normalization methods have significant disadvantages and deficiencies, as discussed in that section of the current application.

The normalization techniques to which the current application is directed are discussed beginning on line 1 of page 17. As stated beginning on line 11 of page 17: "An order-preserving sequence is a sequence of data points in which the values of the data points within the sequence uniformly increase within the sequence. When a sequence is defined as an ordered subset of points within a data set, then a longest-order-preserving sequence ("LOPS") is the maximally sized, one or more ordered subsets of points selected from the data set ***that are ordered by signal strength or by some other associated value, parameter, or characteristic***" (emphasis added). Figures 10A-D of the current application illustrate a two-dimensional LOPS. As stated in the current application in the paragraph that begins on line 23 of page 17:

> In Figure 10A, a number of data points, such as data point 1002, are distributed in a two-dimensional space defined by an orthogonal coordinate system. The positive, horizontal axis 1004 corresponds to a first coordinate x, and the vertical axis 1006 in Figure 10A is the positive axis for the coordinate y. **Each data point, such as data point 1002, has an identify as well as an x value and a y value represented by the position of the data point within the two-dimensional space.** Commonly, the data points are associated with Cartesian coordinates (x,y) where x is the value of the data point with respect

> to the *x* axis and *y* is the value of the data point with respect to the *y* axis. Such a two-dimensional distribution may arise from scanning a microarray at two different frequencies, with the *x* values representing signal intensities scanned at one frequency, and the *y* values representing signal intensities scanned at another frequency. (emphasis added)

As discussed beginning on line 7 of page 18, determination of a LOPS requires that a means of establishing the order of data points must be defined. An exemplary ordering means is defined in that paragraph, with reference to Figure 10B. Construction of a LOPS is illustrated in Figure 10D, and described in the paragraph beginning on line 7 of page 19. In the paragraph beginning on line 26 of page 19 of the current application, a three-dimensional LOPS is discussed with reference to Figures 11A-C.

As discussed in the paragraph beginning on line 28 of page 21 of the current application:

> One embodiment of the present invention is selection of data points within an *n*-dimensional distribution of data points for normalization that coincide with, or fall closely to, a LOPS, or the set of data points coincident with one or more LOPS, within an *n*-dimensional distribution. However, to be practical, there must be a relatively efficient computational method for calculating sets of LOPS points from *n* different data sets. Moreover, because of the increasing constraint represented by selecting points coincident with one or more LOPS as the number of dimensions increases, the computational method needs to be able to systematically relax, to some degree, the LOPS constraints in order to acquire a sufficient number of normalization data points for statistical reliability of the subsequent normalization process.

Note that, as the number of dimensions of a LOPS-construction problem increases, selecting a LOPS from a high-dimensional set of data sets is far from trivial, and requires increasingly complex and time-consuming computations.

One method of the present invention is discussed, beginning on line 9 of page 22, with reference to Figures 12-16D. Graphical and numerical representations of the simple, exemplary data sets are shown in Figures 12 and 13. Figures 14A-K illustrate a technique for computing a LOPS, which uses the two-dimensional table that includes numeric data-set values along with three additional one-dimensional arrays S_{up}, S_{down}, and S_{total}. The computational method discussed with reference to Figures 14A-K is rather complicated, and is best appreciated from the description in the current application. A full C++-like pseudocode implementation of a method for computing LOPS or near-LOPS sets is provided in the current application beginning on line 20 of page 33.

In rejecting the current claims, the Examiner cites a chapter from an elementary calculus and analytical geometry text. The cited reference has absolutely nothing at all to do with data normalization, determination of ordered sets, computation of LOPS or near-LOPS data sets, or anything else even remotely related to the subject matter to which the current application and claims are directed. Those familiar with mathematics and computing would immediately understand that the calculus and analytic geometry text falls within the field of mathematics referred to as "continuous mathematics," while the LOPS-determination methods to which the current application and claims are directed falls within an entirely different field of mathematics, referred to as "discrete mathematics." As discussed in M.P.E.P §2141(III), citing the recent U.S. Supreme Court decision *KSR v. Teleflex*: "'[R]ejections on obviousness cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.'" The current rejections are simply a set of conclusory statements based on a misinterpretation and misstatement of a section of a high-school-level of first-year-undergraduate calculus text. There is no rational underpinning of any kind for these rejections.

Please consider claim 1 as representative of the independent claims of the current invention:

> 1. A method for selecting a set of normalizing data points from *n* data sets, where *n* is at least 3, containing data points having values and identities, the method comprising:
>
> receiving *n* data sets;
>
> considering the data points to be distributed in an *n*-dimensional data-point space;
>
> determining one or more order-preserving sequences of data points within the *n*-dimensional data-point space,
>
> selecting, as normalizing data points, data points from the one or more order-preserving sequences; and
>
> storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets.

The first step recites "receiving *n* data sets." Data sets, as discussed in the current application, comprise data points or data features, described in the current application as follows:

> Data points, or features, in a number of microarray data sets have both identities and values. The values of a data point are generally a measure of scanned intensities of light or radiation emitted from labeled target molecules bound to the feature, and the identity may be two-coordinate indexes, a

> sequence number, or an alphanumeric label that uniquely identifies the feature within the data set. A data point may also, in certain cases, be associated with a weight, where the weight expresses a measure of confidence, constancy, or some other parameter or characteristic.

Appellants' representative has pointed out to the Examiner repeatedly that it is the Appellants' right to define the language used in claims in the text of a patent application, or specification, and that Appellants' definitions control the interpretation of the claims. As explained in the above quote from the current application discussing Figure 10A, data points are associated with both an identity and one or more experimentally determined values. The identity of a data point, for example, is generally a sequence number or an alphanumeric label that uniquely identifies the data point in the data set, such as relating the data point to a feature of a microarray. The data point also has one or more values corresponding to, in the case of microarray data, a signal-intensity value. In the case shown in Figure 10A, the x and y coordinates of data points are the signal intensities associated with the data points that are scanned at different frequencies, the identities of the data points provide a one-to-one mapping of these identities to particular features of a microarray.

Larson has absolutely nothing whatsoever to do with data sets. Larson discusses coordinate systems and vectors, both purely abstract concepts that have nothing to do with data sets and data processing. Larson does not once teach, mention, or suggest anything at all related to data or data points, and does not use the terms "data" and "data set." In the response-to-arguments portion of the Office Action, the Examiner states: "Larson et al. never uses the term 'data.' This is not persuasive because the term data does not need to be recited to fulfill the requirements of the instant claim. Data is any form of information (i.e. the coordinates of a rectangular solid in the first octant of three dimensional space)." Here, the Examiner is attempting to substitute the Examiner's arbitrary and completely unsupported and unsupportable definition of the term "data set" for a very clear and precise definition of that phrase in the current application, quoted above. This violates basic principles of claim interpretation. It is the Appellants, rather than the Examiner, who define the terms and phrases of a claim. Absent such definition, terms and phrases of claims are interpreted based on what those ordinarily skilled in the art would find the terms and phrases to mean. Only when terms and phrases are not defined in the application and are not well defined among those skilled in the art does one then turn to any of various additional, external sources of claim-term and claim-phrase definitions (*see* discussion of *Phillips v. AWH Corp.* throughout

M.P.E.P. § 2111). However, it is never the case claim terms and phrases can arbitrarily be assigned definitions by an examiner. In fact, the Examiner's definition is contrary to the well-understood meaning of the term "data set" in science and mathematics. Coordinates and coordinate systems are not data. Coordinates and coordinate systems are arbitrary conventions used to do describe mathematical abstractions. By contrast, as discussed in the current application and the above-quoted definitions, data points within data sets are associated with experimentally observed values. For this reason alone, it is abundantly clear to anyone cursorily familiar with scientific data processing and data normalization that Larson is completely unrelated to the currently claimed subject matter.

The second element of claim 1 recites "considering the data points to be distributed in an n-dimensional data-point space." The current application provides graphical representations of various different n-dimensional data-point spaces, including a two-dimensional data-point space shown in Figure 10A and a three-dimensional data-point space shown in Figure 11A. As explained in the above-quoted passages of the current application, and in the portions of the specification that refer to Figures 10A-13, a data-point space is not simply some arbitrary collection of points, but, instead, is a distribution of data points from multiple data sets, in which the experimentally observed values associated with data points are mapped to coordinate dimensions. Thus, as discussed above, each data point in Figure 10A has a separate identity, namely a string or number that associates the data point with a particular feature of a microarray, as well as signal intensities measured for two different wave lengths that are mapped to the x and y coordinates.

Larson, by contrast, is concerned with abstract mathematical concepts, such as coordinate systems and vectors. The points in Cartesian three-dimensional space are not data points. Points in Cartesian three-dimensional space are entirely and completely defined by their coordinates. They are not associated with any experimentally observed value and are not distributed in any kind of data-point space.

The Examiner states:

> Now, when taking into account limitations of instant claim 1 in combination with the rectangular solid (i.e. Figures 14.1 and 14.6); a rectangular solid in the first octant has 8 data points with the following identifiers → (0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), and (1,1,1). Consequently, these coordinate are in three dimensional space and they encompass three data sets: X coordinates [0,1,0,0,1,1,0,1], Y coordinates [0,0,1,0,1,0,1,1], and Z coordinates[0,0,0,1,0,1,1,1]. Consequently, the Figures of Larson illustrate three data sets in three dimensional space.

Figure 14.1 of Larson, a simple, standard representation of the positive portion or positive octant of three-dimensional Cartesian space. There is no rectangular solid in Figure 14.1. Larson explicitly describes Figure 14.1 as showing "the positive portion of each coordinate axis." Even were there a rectangular solid, rectangular solids do not have data points. Rectangular solids are continuous volumes. Figure 14.6 shows a vector. Vectors are not rectangular solids. The three ordered triples in Figure 14.6 (0,0,1), (0,1,0), and (1, 0, 0) denote the unit vectors **i**, **j**, and **k**, as explicitly stated by Larson in the text referring to Figure 14.6. The other ordered triples recited by the Examiner occur in neither figure. Appellants' representative has no idea to what the Examiner is referring by the phrase: " X coordinates [0,1,0,0,1,1,0,1], Y coordinates [0,0,1,0,1,0,1,1], and Z coordinates[0,0,0,1,0,1,1,1]." These sets have no meaning with respect to anything in Larson, and are not shown or discussed in Larson. The Examiner apparently thinks that a set of coordinates with respect to coordinate axis constitutes a data set. This, of course, completely contradicts the definition of data sets and data points, quoted above, in the current application. It is impossible for the points identified by the Examiner in three-dimensional Cartesian space to be data points according to the definition of the phrase "data point" in the current application, since each such point requires three coordinate values to be identified, and there are no other numbers or values associated with the points. Those even cursorily familiar with modern mathematics well understand that a point in three-dimensional Cartesian coordinate space is simply an abstraction. Although data points may, indeed, be mapped to a coordinate space, such mappings require mapping of data values associated with data points to coordinates. Of course, in Larson, there is no such mapping, because Larson is not concerned with data points, but instead is merely concerned with points and vectors in three-dimensional Cartesian coordinate space, which are purely mathematical abstractions. Again, the word "data" and the phrases "data set," "data-point space," and other such terms defined in the current application do not occur even once in Larson.

The third step of claim 1 recites "determining one or more order-preserving sequences of data points within the *n*-dimensional data-point space." In the current application, order-preserving sequences of data points are introduced in the following passage that begins on line 11 of page 17:

> An order-preserving sequence is a sequence of data points in which the values of the data points within the sequence uniformly increase within the sequence. When a sequence is defined as an ordered subset of points within a data set,

> then a longest-order-preserving sequence ("LOPS") is the maximally sized, one or more ordered subsets of points selected from the data set that are ordered by signal strength or by some other associated value, parameter, or characteristic.

Beginning on line 7 of the current application on page 18, the application further states:

> In order to determine a LOPS for a two-dimensional distribution of data points, such as that shown in Figure 10A, a means of establishing the order of data points must be first defined.

Those familiar with mathematics well understand that an order-preserving sequence does not simply naturally arise from a data set, but instead involves first defining a method for establishing an order of data points, and then applying the method to the data points within a data set in order to select data points that together comprise a LOPS. In the two- and three-dimensional cases shown in the current application, for ease and clarity of illustration, ordering of data points can be defined according to a geometrically intuitive ordering rule. However, in higher-dimensional data-point spaces, including that shown in Figure 12, the task becomes more conceptually difficult.

The Examiner states:

> Now, the next step of determining one or more order preserving sequence in 3-dimensional space is illustrated in Figure 14.6 on page 787 of Larson et al., wherein it is shown that a unit vector from [0,0,0] to [1,1,1] preserves the increasing order of each data set (x coordinates, y coordinates and z coordinates). Consequently, Figure 14.6 also illustrates selection of the sequence of points [0,0,0] and [1,1,1].

The Examiner is reading the step of determining order-preserving sequences onto Figure 14.6 of Larson. Figure 14.6 of Larson shows the three standard unit vectors in Cartesian three-dimensional space, generally referred to as **i**, **j**, and **k**. A unit vector has a length of 1, as is well known to those with high-school mathematics training. Those with high-school-level mathematics training can appreciate that there are no ordered sequences shown or implied in Figure 14.6 of Larson. The Examiner's statements make absolutely no sense. First of all, a vector from point (0,0,0) to (1,1,1) is not a unit vector, as is well understood by those familiar with high-school mathematics. Indeed, in the following page of Larson, the definition of the length of a vector is provided, and, according to that definition, and as well understood by anyone familiar with mathematics, the length of a vector from (0,0,0) to (1,1,1) is $\sqrt{3}$; which is not equal to 1, whereas a unit vector has, by definition, a length equal to 1. The Examiner has clearly failed to understand the subject matter represented by Figure 14.6 in Larson.

Second, the Examiner has apparently failed to understand the definition of LOPS in the current application. An order-preserving sequence preserves the order of experimentally observed values associated with data points. When experimentally observed values are mapped to coordinates of a coordinate space, as is done in the data-point sets illustrated in Figures 10A and 11A of the current application, then there may well be an intuitive, geometrical approach to determining the relative order of two data points. However, Larson is not in any way concerned with ordering anything, including data points, or even Cartesian points, in Figure 14.6. There is no ordering whatsoever implied by the unit vectors **i**, **j**, and **k**. All three unit vectors have the identical length of 1, and the unit vectors are orthogonal to one another, and are thus independent in the familiar vector-space or linear-algebra notion of "independence." Unit vectors are, in fact, in many ways quintessentially unordered. There is no obvious or reasonable metric or criteria for assigning a relative order between two unit vectors. Direction is certainly not an ordering criterion, nor obviously is length, since all unit vectors have the same length. A single vector, as is well understood by those familiar even with high-school mathematics, is defined by a point, a length, and a direction. A unit vector is defined by a point and a direction. A vector does not encompass or include a set of points. It is a purely abstract mathematical quantity defined by, in the case of unit vectors, one point and one direction, with the direction commonly specified by two angles or three direction cosines. A vector does not and cannot preserve any kind of increasing order of coordinates, as asserted by the Examiner.

There is no definition or even mention of ordering or sequences in Figure 14.6 of Larson, or anywhere in the cited pages of Larson. As clearly discussed in the current application, the first step in determining a LOPS is to establish some means, metric, or definition for ordering data points. As discussed at length in the current application, a LOPS comprises a set of points selected from a data set. Larson does not teach, mention, or suggest selection of data points from any kind of data set, and does not even teach, mention, or suggest selection of particular points from three-dimensional Cartesian space. It is true that one could define a directed line segment as a subset of the points in three-dimensional Cartesian space, and define an order of the points of the line segment in terms of coordinates of each point. For certain line segments with certain directions, all of the coordinates of all of the points along the line segment would increase uniformly. However, in other cases, certain of the coordinates would increase while other of the coordinates would decrease or remain constant. Whether or not a line segment abstractly represents an ordered set of points

critically depends on the definition of the ordering criteria, and Larson obviously does not teach, mention, or even remotely suggest any such ordering criteria.

Larson involves continuous mathematics. There are an infinite number of points, which are mathematical abstractions, along any line segment, regardless of its length. The current application is, by contrast, directed to discrete mathematics, in which there is a discrete number of data points within a data set. Claim 1 discusses finite data sets and finite LOPS. The Examiner appears to be attempting to draw analogies to continuous, infinite sets in Larson's three-dimensional Cartesian coordinate space. An infinite set of points cannot be stored in any computer memory. As those familiar with computer science and electronics well understand that computer memories are finite.

In summary, Larson has nothing at all do with data sets, data points, order-preserving sequences, or selection of data points or any other kind of points from data sets and order-preserving sequences. Instead, Larson is a simple, high-school or first-year-undergraduate calculus text with figures that show rudimentary mathematical abstractions, such as the positive octant of three-dimensional Cartesian coordinate space and the definition of the unit vectors **i**, **j**, and **k**. Neither abstraction involves or in any way teaches, mentions, or suggests any kind of ordering or sequence. The Examiner is choosing various geometrical points from these figures and claiming that, by doing so, he has found order-preserving sequences. He has not. Furthermore, the Examiner fails to appreciate that it is the cited reference, rather than the Examiner, which needs to teach, mention, or suggest the selection and ordering process. Larson, obviously, fails to do so. Had the Examiner successfully defined data points and data sets and mapped those to Larson's figures, and then selected an order-preserving sequence of points from those mapped to the figures, the Examiner may well have created an order-preserving sequence. However, that would have had nothing whatsoever to do with whether or not Larson teaches, mentions, or suggests such order-preserving sequences. Clearly, Larson does not. Whether or not the Examiner can use the teachings of the current application to create data sets, map data points to geometrical points in Cartesian space, define ordering metrics and relations for the points, and construct order-preserving sequences has absolutely nothing whatsoever to do with whether Larson teaches, mentions, or suggests such operations.

CONCLUSION

The Examiner has failed to cite any rule, decision, or statute for the propositions under which claims 1-12 and 14-28 have been rejected as being directed to non-statutory subject matter. The rejections are therefore improper and unfounded. Larson, the only reference cited by the Examiner in rejecting claims 1, 4, 7, 12, 14, 17, 20, 23, and 28 under 35 U.S.C. §103(a), is entirely unrelated to the current application and current claims, and discusses a completely different field of mathematics than the field of mathematics to which the currently claimed methods and systems are related. Larson fails to teach, mention, or suggest any of the steps or elements of any of the current claims, and fails to teach, mention, or suggest any of the Examiner's conclusions and interpretations of Larson's Figures 14.1 and 14.6. In fact, the Examiner has rather completely failed to understand those Figures, and how they are related to the text in Larson that refers to them. Clearly, the Examiner has not even remotely made a *prima facie* case of an obviousness-type rejection, by any of the standards in M.P.E.P. §2142. There is simply no rational underpinning for any of the Examiner's attempts to read current claim language onto Figures 14.1 and 14.6 of Larson.

Appellants respectfully submit that all statutory requirements are met and that the present application is allowable over all the references of record. Therefore, Appellants respectfully request that the present application be passed to issue.

Respectfully submitted,
Zohar Yakhini et al.
OLYMPIC PATENT WORKS PLLC

By
Robert W. Bergstrom
Registration No. 39,906

Olympic Patent Works PLLC
P.O. Box 4277
Seattle, WA 98104
206.621.1933 telephone
206.621.5302 fax

CLAIMS APPENDIX

1. A method for selecting a set of normalizing data points from *n* data sets, where *n* is at least 3, containing data points having values and identities, the method comprising:

receiving *n* data sets;

considering the data points to be distributed in an *n*-dimensional data-point space;

determining one or more order-preserving sequences of data points within the *n*-dimensional data-point space,

selecting, as normalizing data points, data points from the one or more order-preserving sequences; and

storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets.

2. The method of claim 1 wherein the one or more order-preserving sequences of data points is a single, longest order-preserving sequence of data points.

3. The method of claim 1 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights.

4. The method of claim 1 wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence.

5. The method of claim 1 wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points of lengths within a threshold value of the length of an order-preserving sequence of data points of maximum length.

6. The method of claim 1 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points with sums of weights within a threshold value of the sum of weights of an order-preserving sequence of data points with a greatest sum of weights.

7. The method of claim 1 wherein considering the data points to be distributed in an n-dimensional data-point space further includes, for each data point, considering the data point to have a value in each of n-dimensions, the value of a data-point in an ith dimension equal to the value of the data point in an ith data set, where $1 \le i \le n$.

8. The method of claim 1 wherein determining an order-preserving sequence of data points within the n-dimensional data-point space further includes:

for each currently considered dimension,

ordering the data points with respect to the currently considered dimension;

traversing the ordered data points in a first direction, determining a metric corresponding to a maximum subsequence for each data point in the first direction; and

traversing the ordered data points in a second direction, determining a metric corresponding to a maximum subsequence for each data point in the second direction;

summing the determined metrics for each data point in each dimension to produce a metric sum for each data point; and

selecting as belonging to the maximum order-preserving sequence of data points those data points having a greatest metric sum.

9. The method of claim 8 wherein selecting, as normalizing data points, data points from the order-preserving sequence further includes selecting data points with a metric sum greater than a threshold value.

10. The method of claim 8 wherein selecting, as normalizing data points, data points from the one or more order-preserving sequences further includes selecting data points of a single order-preserving sequence.

11. The method of claim 8 wherein selecting, as normalizing data points, data points from the one or more order-preserving sequences further includes selecting data points that most evenly partition the data points into subsets of data points.

12. Computer instructions stored in a computer readable medium that implement the method of claim 1.

13. Cancelled

14. A system for selecting a set of normalizing data points from *n* data sets, where *n* is at least 3, containing data points having values and identities, the system comprising:

a processor;

a memory;

and computer instructions that select the set of normalizing data points from *n* data sets by

receiving *n* data sets,

considering the data points to be distributed in an *n*-dimensional data-point space,

determining one or more order-preserving sequence of data points within the *n*-dimensional data-point space, and

selecting, as normalizing data points, data points from the one or more order-preserving sequences; and

storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets.

15. The system of claim 14 wherein the one or more order-preserving sequences of data points is a single, longest order-preserving sequence of data points.

16. The-system of claim 14 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights.

17. The system of claim 14 wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence.

18. The system of claim 14 wherein the one or more order-preserving sequences of data points are order-preserving sequence of data points within a threshold value of an order-

preserving sequences of data points of maximum length.

19. The system of claim 14 wherein the one or more order-preserving sequences of data points are order-preserving sequence of data points within a threshold value of an order-preserving sequences of data points with a greatest sum of weights.

20. A method for selecting a set of normalizing data points from *n* data sets, where *n* is at least 4 and even, containing data points having values and identities, the method comprising:

receiving *n* data sets;

considering the data points to be distributed in $\frac{n}{2}$ 2-dimensional data-point spaces;

determining one or more order-preserving sequences of data points for each of the $\frac{n}{2}$ 2-dimensional data-point spaces~~; and~~,

selecting, as normalizing data points, data points from the order-preserving sequences; and

storing the selected normalizing points in a computer memory as a basis for subsequent normalization of the *n* data sets.

21. The method of claim 20 wherein the one or more order-preserving sequences of data points is a single, longest order-preserving sequence of data points.

22. The method of claim 20 wherein the data points within *n* data sets are associated with weights and wherein the one or more order-preserving sequences of data points is an order-preserving sequence of data points with a greatest sum of weights.

23. The method of claim 20 wherein the one or more order-preserving sequences of data points is a longest order-preserving sequence of data points having a shortest Euclidian distance accumulated along a path from an initial data point of the order-preserving sequence to a final data point of the order-preserving sequence.

24. The method of claim 20 wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points within a threshold value of an order-

preserving sequence of data points of maximum length.

25. The method of claim 20 wherein the data points within n data sets are associated with weights and wherein the one or more order-preserving sequences of data points are order-preserving sequences of data points with sums of weights within a threshold value of the sum of weights of an order-preserving sequence of data points with a greatest sum of weights.

26. The method of claim 20 wherein determining an order-preserving sequence of data points within a 2-dimensional data-point space further includes:

for each currently considered dimension,

ordering the data points with respect to the currently considered dimension;

traversing the ordered data points in a first direction, determining a metric corresponding to a maximum subsequence for each data point in the first direction; and

traversing the ordered data points in a second direction, determining a metric corresponding to a maximum subsequence for each data point in the second direction;

summing the determined metrics for each data point in each dimension to produce a metric sum for each data point; and

selecting as belonging to the maximum order-preserving sequence of data points those data points having a greatest metric sum.

27. The method of claim 20 wherein selecting, as normalizing data points, data points from the one or more order-preserving sequences further includes selecting data points which occur in the one or order-preserving sequences computed for greater than a threshold fraction of the $\frac{n}{2}$ 2-dimensional data-point spaces.

28. Computer instructions stored in a computer readable medium that implement the method of claim 20.

29. Cancelled

EVIDENCE APPENDIX

None.

RELATED PROCEEDINGS APPENDIX

None.

AGILENT TECHNOLOGIES, INC.
Legal Department, DL429
Intellectual Property Administration
P. O. Box 7599
Loveland, Colorado 80537-0599

ATTORNEY DOCKET NO. 10020708-1

OIPE IAP35 NOV 14 2008 PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Zohar Yakhini et al.

Serial No.: 10/825,893

Examiner: Russell Scott Negin

Filing Date: April 16, 2004

Group Art Unit: 1631

Title: METHOD AND SYSTEM FOR NORMALIZATION OF MICROARRAY DATA

COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria VA 22313-1450

TRANSMITTAL LETTER FOR RESPONSE/AMENDMENT

Sir:

Transmitted herewith is/are the following in the above-identified application:

☐ Response/Amendment ☐ Petition to extend time to respond

☐ New fee as calculated below ☐ Supplemental Declaration

☐ No additional fee (Address envelope to "Mail Stop Amendments")

☒ Other: Response to Non-Compliant Appeal Brief (Fee $_____)

CLAIMS AS AMENDED BY OTHER THAN A SMALL ENTITY						
(1) FOR	(2) CLAIMS REMAINING AFTER AMENDMENT	(3) NUMBER EXTRA	(4) HIGHEST NUMBER PREVIOUSLY PAID FOR	(5) PRESENT EXTRA	(6) RATE	(7) ADDITIONAL FEES
TOTAL CLAIMS		MINUS		= 0	X 50	$ 0
INDEP. CLAIMS		MINUS		= 0	X 200	$ 0
☐ FIRST PRESENTATION OF A MULTIPLE DEPENDENT CLAIM					+ 360	$ 0
EXTENSION FEE	1ST MONTH 120.00 ☐	2ND MONTH 450.00 ☐	3RD MONTH 1020.00 ☐	4TH MONTH 1590.00 ☐		$ 0
					OTHER FEES	$ 0
					TOTAL ADDITIONAL FEE FOR THIS AMENDMENT	$ 0

Charge $0 to Deposit Account **50-1078**. At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account **50-1078** pursuant to 37 CFR 1.2 5. Additionally please charge any fees to Deposit Account **50-1078** under 37 CFR 1.16, 1.17, 1.19, 1.20 and 1.21. A duplicate copy of this transmittal letter is enclosed.

Respectfully submitted,

Zohar Yakhini et al.

By Robert W. Bergstrom

Robert W. Bergstrom
Attorney/Agent for Applicant(s)

Reg. No. 39,906

Date: November 12, 2008

Telephone No. 206.621.1933

I hereby certify that this correspondence is being Deposited with the United States Postal Service as First class mail in an envelope addressed to: Commissioner for Patents, PO Box 1450, Alexandria, VA 22313-1450.

Date of Deposit: November 12, 2008

Typed Name: Joanne Bourguignon

Signature: Joanne Bourguignon